Enhancing the data usability from the ECOTOX Knowledgebase

Colleen Elonen*, Jennifer H. Olker*, Stephen Erickson†, Michael Skopinski†, Anne Pilli†, Brian Kinziger†, Kathryn Murphy†, Gage Sachs†, and Dale Hoff*

*U.S. EPA, ORD/CCTE/GLTED, Duluth, MN; †General Dynamics Information Technology, Duluth, MN

www.epa.gov

Colleen Elonen | Elonen.colleen@epa.gov | 218-529-5195

What is the ECOTOX Knowledgebase?

The ECOTOX Knowledgebase (ECOTOX) is a comprehensive, publicly available application providing chemical and environmental toxicity data on aquatic life, terrestrial plants and wildlife. Data are curated from the scientific literature after exhaustive search protocols and contains more than 1 million data records from over 12,000 chemicals and 13,000 species.

Over the last 30+ years we have continually strived to improve the searchability and functionality of the Knowledgebase. Here we present refinements of ECOTOX's current tools and enhancements to improve the usability of the Knowledgebase to different database tools and data platforms.

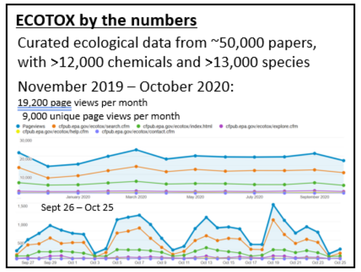
ECOTOX by the numbers

Curated ecological data from ~50,000 papers, with >12,000 chemicals and >13,000 species

November 2019 – October 2020:

19,200 page views per month

9,000 unique page views per month



Applicable information to support study evaluations

ECOTOX conducts comprehensive literature searches and reviews of toxicity data in open and grey literature with transparent standard operating procedures that meet the requirements of systematic review protocols. While this approach streamlines the cost for literature searches and data curation within the Agency, it also provides all applicable test data information in a format that can be readily used for study evaluations by government or public entities.

Criteria for inclusion in ECOTOX

Table 1. Criteria and requirements for inclusion in ECOTOX following well-established Standard Operating Procedures, compared to the recently-developed PECO statement.

PECO statement for ECOTOX		Requirements/Inclusionary Criteria from ECOTOX SOP
P (Population)	Animal: Aquatic and terrestrial species (live, whole organism) of any lifestage (including preconception, in utero, lactation, peripubertal, and adult stages). Include wild mammals (e.g.: Peromyscus sp.), insects, spiders, amphibians, birds, crustaceans, fish, molluscs, reptiles, worms and invertebrates. Bacteria and viruses are not included. Plants: Aquatic and terrestrial species (live), all plants including algal, moss, lichen and fungi species.	• Ecologically-relevant species • Live, whole organisms • Organism taxonomic information verifiable against standard taxonomic sources • Priority species are wild (test results for terrestrial domestic and laboratory species are used to fill data gaps when needed) • In vitro studies (with viable cells or tissue) flagged for possible inclusion as requested by Programs • NOT: humans, monkeys, bacteria, viruses, yeast
E (Exposure)	Relevant forms: Chemical of Concern, name and CASRN (plus synonyms, tradenames); when requested: Metabolites, degradants, parent compound and related chemicals. Animal: Any exposure to relevant forms of the chemical of concern including via water, injection, diet, and dermal, with reported concentration and duration. Inhalation studies are excluded unless this is the primary route of environmental exposure (e.g., for volatile compounds). Plants: Exposure to relevant forms of the chemical of concern via water or soil, with reported concentration and duration. * Studies involving exposures to mixtures will be included only if they include exposure to a relevant form for the chemical alone. * Chemical exposures for aquatic organisms where only sediment concentrations are reported from field studies are excluded (unless porewater concentration measured); laboratory-based sediment studies are retained.	• Verifiable Chemical Abstract Services (CAS) number • Single chemical exposure • Relevant to environmental exposure • Report exposure concentration, dose or application rate • Report duration of exposure • Sediment studies must have a water concentration reported to be included • NOT: Air pollution studies related to CO2 and ozone
C (Comparison / Control)	A concurrent control group exposed to vehicle-only treatment and/or untreated control (control could be a baseline measurement).	• Must have a control treatment
O (Outcome)	All biological effects (including bioaccumulation from laboratory studies with concurrently measured water and tissue concentrations).	• Biological effect measured • Effect concurrent with associated chemical exposure • Adverse effects are priority (beneficial, nutritional effects are lower priority)
Publication/ Data Format		• Primary source of the data • Study must be a full article in English • NOT: Reviews or abstract only

Data extraction fields in ECOTOX

Table 2. Types of data extracted from each reference (if applicable and reported) with sample data fields by category, and examples of how ECOTOX fields can inform study evaluation questions.

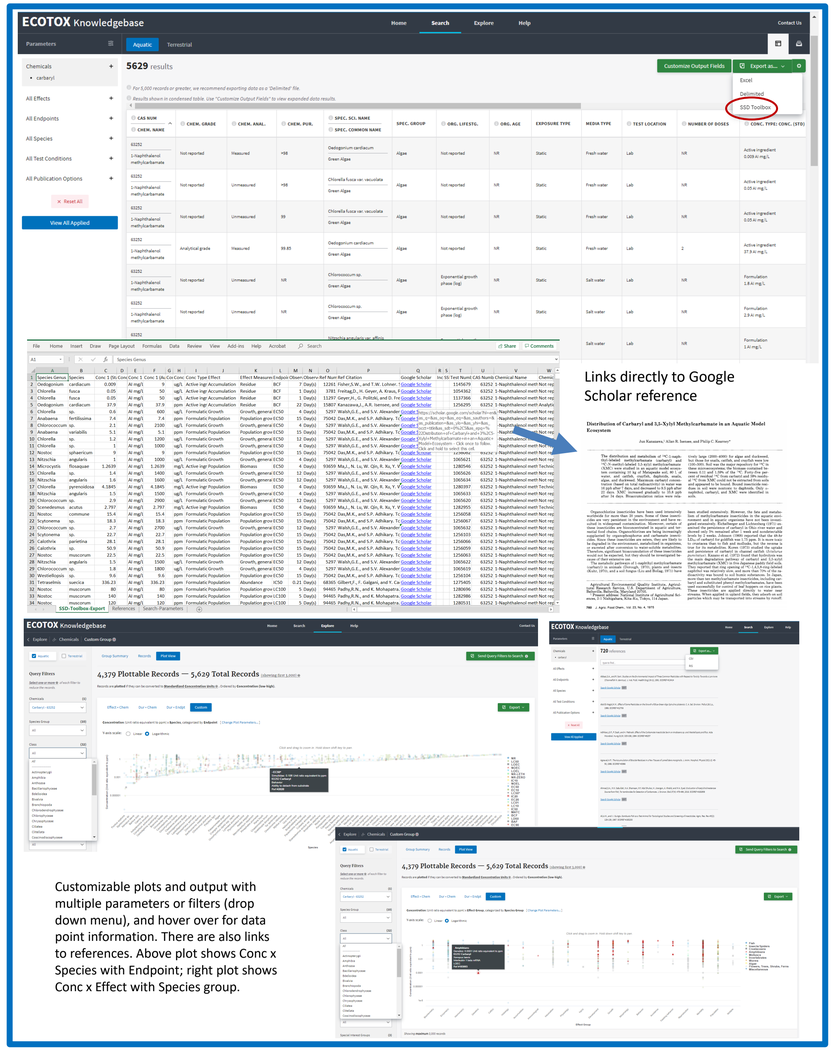
Category	Data Fields (not all inclusive)	Select study evaluation questions with relevant ECOTOX field(s)
Chemical	Chemical Name, CASRN, Grade, Purity, Formulation, Carrier *Test Specific:* Analysis, Application Type and Rate/Frequency, Number of Doses, Doses, Concentration Type (e.g., active ingredient or formulation), Concentration/Dose associated with each effect and/or endpoint	Is test substance identified? **Required for inclusion in ECOTOX inclusion** Is the purity of test substance reported? Chemical Purity Were chemical concentrations verified? Chemical Analysis (e.g., nominal versus measured concentrations)
Species	Scientific and Common Name, Taxonomy, Lifestage, Age, Initial and Final Weight, Gender, Source	Is the species given? **Verifiable species (Scientific Name, etc.) required for inclusion in ECOTOX** Are the organisms well described? Organism Source, Lifestage, Age, Gender, Initial and Final Weight
Test Conditions	Test Method, Media Type, Test Location, Exposure and Study Duration, Control, Experimental Design, Physical and Chemical Soil and Water Parameters	Are appropriate controls performed? **A control is required for inclusion in ECOTOX**, type described in Control Is a guideline method (e.g., OECD) used? Test Method Are the experimental conditions appropriate and acceptable for the test substance and organism? Test Method, Media Type, Test Location, Experimental Design, Physical and Chemical Soil and Water Parameters (e.g., pH, Temperature, Dissolved Oxygen)
Test Results	Effect (observation of a response): general effect groups and specific effect measurements, Endpoint (quantification of an observed effect, e.g., LC50), Trend, Response Site, Effect %, Statistical Significance and Level, Observed Duration (exposure Duration when result observed), Bioconcentration (BCF or BAF) with units	Are the reported effects and endpoints appropriate for the purpose, test substance and organism? Effect Measurement, Endpoint Is the response/effect statistically significant? Statistical Significance, Significance Level

U.S. Environmental Protection Agency
Office of Research and Development

The views expressed in this poster are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA.

Updated data filtering, visualization and graphing options

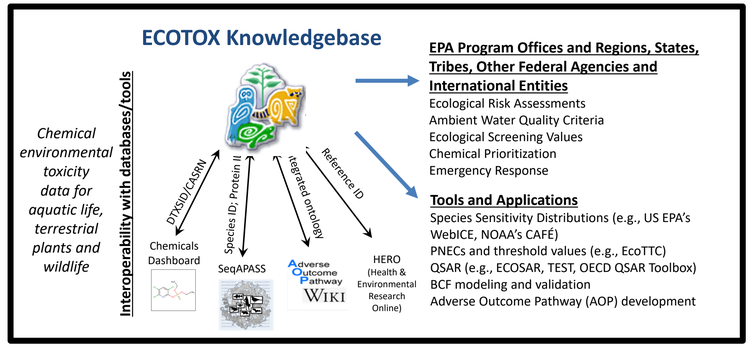
Customizable outputs were developed to allow users to identify key information and select data fields for export into formats compatible with direct input into platforms such as EPA's SSD Toolbox and populating data evaluation records.



For further information on the ECOTOX Knowledgebase, contact ECOTOX Support: ecotox.support@epa.gov
ECOTOX originated in the early 1980s and is maintained by U.S. EPA ORD, available at: www.epa.gov/ecotox

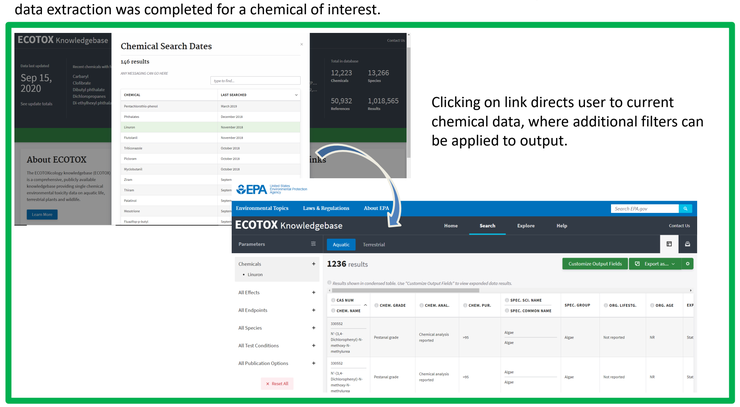
Updated interoperability and linkages

Efforts have been made to increase interoperability and integration with other platforms such as EPA's Chemicals Dashboard (which provides a myriad of chemical characteristics) and to SeqAPASS to aid in species extrapolations of chemical effects. Data from ECOTOX are used as a component of many tools and applications, including tools for Species Sensitivity Distributions, Quantitative Structure Activity Relationships, and for modeling and development of Adverse Outcome Pathways.



Addition of Chemical Search history

Information on full chemical search dates, and direct links to current chemical data is now available. This feature allows risk assessors and researchers ability to search the last time a complete literature search and data extraction was completed for a chemical of interest.



Other Upcoming Enhancements

Activities are underway to improve:

Efficiencies in the systematic review/data curation processes, including computer-assisted identification of search terms, automation of literature searches, machine learning/artificial intelligence for identifying applicable studies, and automated data extraction. See presentations 5.17.01 (Hoff) and 5.17.04 (Howard).

Reusability of literature search results and extracted data, including centralized strategy for literature searches, identification of studies vetted for regulatory use, and development of study quality evaluation outputs.

Interoperability with other databases and tools through mapping ECOTOX terms to existing ontologies and development of web services.